

ASSIGNMENT INFORMATION

Due Date

Saturday, October 23, 2021

11:59 PM

Points Possible

1

View Rubric

Construction of a Regression algorithm

The .mat and the .csv files provided is a dataset that contains a training and a test sets. Each set consists of the predictor input X_{train} and X_{test} for regressors y_{train} and y_{test} , corresponding to the response y of a physical model to a vector signal x of 19 dimensions. The train and test outputs consist of 81 samples and they are depicted in the figure below.

In this assignment, you must construct a linear ridge regression that uses inputs X to predict y . Use a 20% of the training data to validate y . The validation procedure consists of:

- Training the predictor with 80% of the training data and a given value for γ .
- Running a test for the rest of the training data.
- Computing the mean square error of the prediction.
- Repeat for a reasonable range of γ .

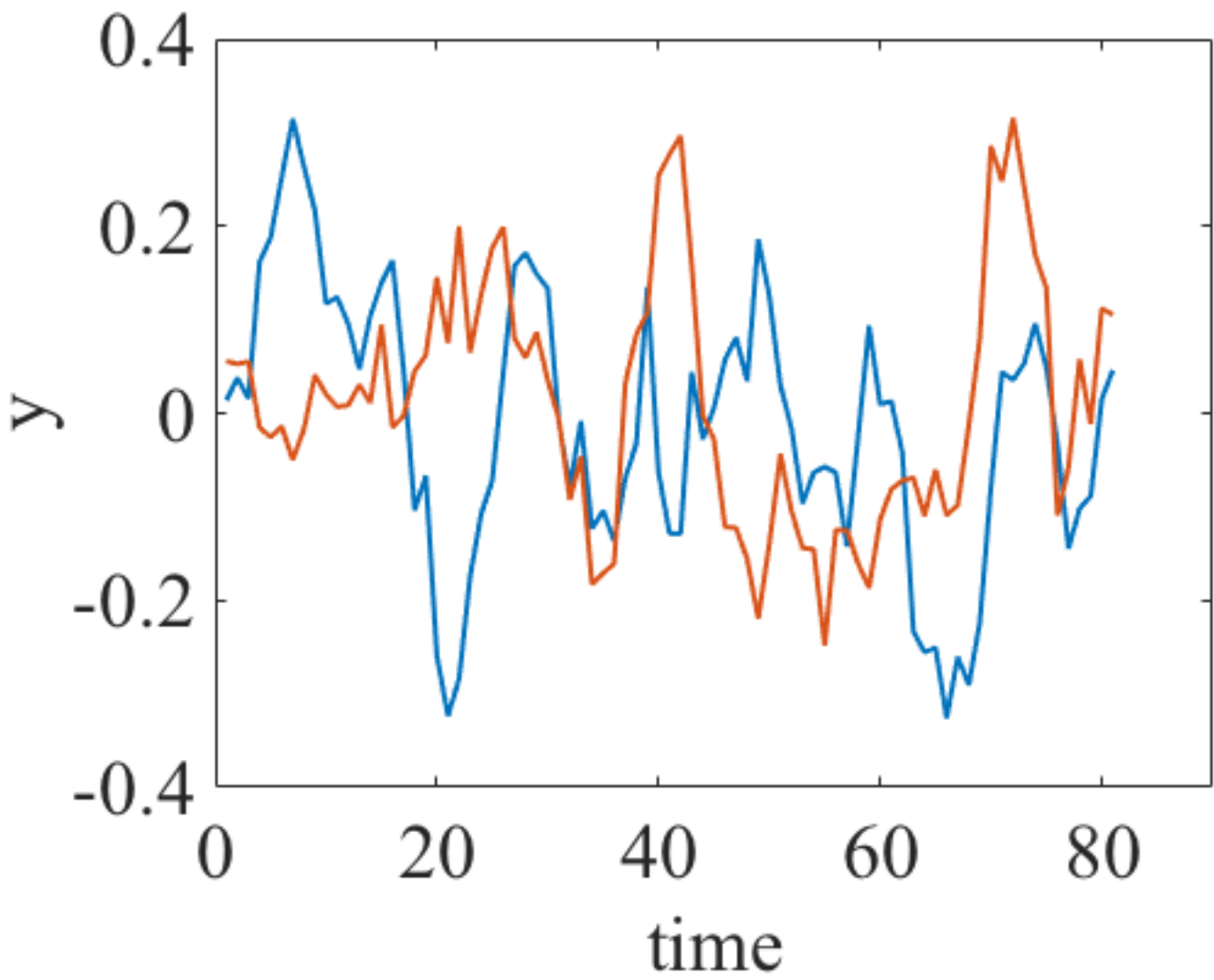
Choose the value of γ that produced the best result. A reasonable interval can be between 0.01 times and 10 times the trace of the matrix, using a logarithmic spacing. In matlab, use function logspace.

Provide the following results of the experiment:

- A graph comparing the real and predicted data.
- A graph of the result of the validation square error.
- The value of the optimal validation and test square errors.

Repeat the experiment above, but using an SVR and a nu-SVR. In both cases you should use a 20% of the training data in order to validate the parameters of the SVM.

- Provide the following results of the experiment:
- A graph comparing the real and predicted data.
- A graph of the result of the validation square error.
- The value of the optimal validation and test square errors.
- A written comparison of the results of this experiment and the previous ones.



NOTE: The csv file is for Python users. The structure of the data is:

- Column 1: ftest (test regressors without noise)
- Column 2: ftrain (train regressors without noise)
- Columns 3-21: Xtest
- Columns 22-40: Xtrain
- Column 41: ytest
- Colun 42: ytrain

[data.csv](#)
[data5.mat](#)
[Validation fo ridge regression](#)